Process for finding likely pathways using machine reading:

1. Write one or more literature queries for use with PubMed
2. Configuration: (/users/../.config/config.ini) INDRA\_DB\_REST\_URL = <https://db.indra.bio> – see INDRA documentation for more details
3. Retrieve PMCIDs for literature corpus
   1. Export results as a list of PMIDs (display 200 per page and filter out any abstracts for fastest turnaround)

Graphical user interface, text, application

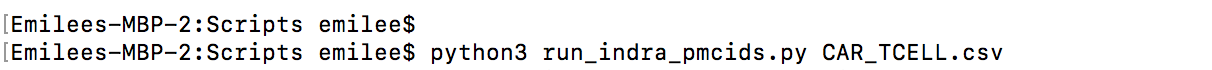
Description automatically generated

* 1. Convert to PMCIDS using the [NIH ID Converter](https://www.ncbi.nlm.nih.gov/pmc/tools/idconv/), and export as a .CSV
     1. Potential pitfalls: some papers will have PMIDs but not PMCIDs

Graphical user interface, text, application, email

Description automatically generated

1. Use script to search the INDRA database for all statements from literature corpus
   1. Run script and input one argument – the name of the CSV file of PMCIDs



* 1. The script will output all formatted statements in a .txt, and also print to screen the ids of all papers that had no statements.
     1. Potential pitfall – the INDRA database may not contain reading or database results from new papers. For the papers that are missing from the INDRA database, it may be possible to process them separately through a reader like REACH, Eidos, etc.